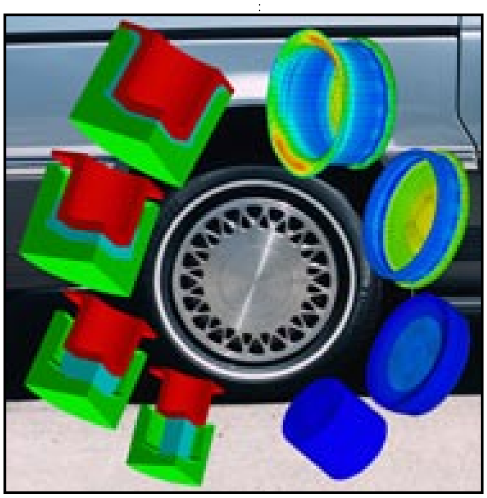
Computations

Today, a large part of research and development at the Laboratory combines computer simulation and experimentation, reducing the need for expensive experiments.

The DYNA3D computer program is one of our most successful industrial partnerships. The LLNL-developed code to analyze impacts upon structures helps American industry maintain a competitive position in the world market. Here the codes are depicted as they apply to automobile wheel components.



Computers have enabled us to extend our design and predictive capabilities and therefore our ability to make breakthrough discoveries in virtually every field of science—from physics, materials science, and chemistry to engineering, geology, biology, and medicine. Among our activities this year, we:

- Created the Center for Computational Science and Engineering to provide a focal point for research in computational science.
- Launched the Industrial Computing Initiative, a collaboration of LLNL, Los Alamos National Laboratory (LANL), and an industry group of 16 organizations headed by Cray Research, Inc.
- Participated in the DOE's consolidated/collaborating SuperLab, an effort to link LLNL, LANL, and the two Sandia labs into a "virtual" laboratory.
- Integrated a massively parallel computer into our production work flow and began extensive conversion of supercomputer codes to parallel architecture.
- Began upgrading the communications capability of the Energy Sciences Network.
- Expanded our collaborations with other laboratories and private industry through additional Cooperative Research and Development Agreements (CRADAs).

The following represent accomplishments in Computations areas for the past year.

Center for Computational Sciences and Engineering Established

The Center for Computational Sciences and Engineering (CCSE) was established to develop and apply advanced computational methodologies to solve large-scale scientific and engineering problems arising in DOE mission areas involving energy, environmental, and industrial technology. The Center's staff pursues a wide range of research activities in areas including computational fluid dynamics, flow in porous media, particle and chemical kinetics, plasma modeling, and general numerical methods for solving large-scale differential/algebraic systems problems.

Technology transfer and industrial collaborations form a major component of Center research activities. For example, we are working with collaborators at LANL, UC Berkeley, and the Courant Institute at New York University to develop sophisticated techniques for modeling industrial combustion problems—techniques that an engineer could use to understand how pollutants such as NOx are formed in an industrial burner.

Researchers from CCSE and other LLNL areas are working with IT Corporation under a CRADA to deploy massively parallel computers for modeling flow in heterogeneous porous media. Detailed simulations of large sites will play a key role in the design of remediation techniques to clean up contaminated groundwater at sites throughout the world. We also initiated a collaboration with a large oil and gas service company to develop a massively parallel simulator for the design of nuclear well-logging tools (probes used to determine lithology and fluid characteristics in the regions around boreholes).

The Industrial Computing Initiative

The Industrial Computing Initiative represents a three-year collaboration between LLNL and LANL and 16 industrial participants to demonstrate that economical solutions to grand-challenge-scale computational problems can provide a competitive advantage in the marketplace. Multiple CRADAs are used to define the relationships and the deliverables among the participants. We intend to develop applications whose added dimensionality and enhanced realism and accuracy create a distinct economic advantage for the companies involved. With each laboratory using a 128-processor Cray Research T3D MP computer, the various projects are targeted at developing either efficient application codes or software that facilitates the use of parallel computers.

The SuperLab Project

The long-term objective of the SuperLab program is to develop the software infrastructure for linking LLNL, LANL, and the two SNL laboratories by means of high-speed digital communication into a single "virtual laboratory." Scientists and engineers at the four sites then will be able to collaborate more closely and more easily share their computing resources, secure as well as unclassified. Many of the technical components of the SuperLab are strongly synergistic in discharging the Defense Program's missions. For example, an "intelligent archive" project would be useful in a full range of applications, from environmental data structures to weapons-design data-legacy files.

Massively Parallel Computing Integrated into Production Workloads

As we continue to move from conventional supercomputing to a massively parallel environment, the Livermore Computer Center is using its Meiko CS-2 MPP (more than 64 of its 256 processors are now operating) and a migration strategy called the Livermore Model, which enables us to partition a massively parallel environment so that monoprocessing, multiple processing, and massively parallel processing can be performed simultaneously. Since we are just beginning to convert our key applications to massively parallel operation, we are initially running a large number of serial jobs simultaneously on separate nodes (that is, running them in a capacity mode). We are migrating them to parallel operation one package at a time, starting with those that will benefit most from the increased capability (that is, the simultaneous focusing of many nodes on a single parallel application). For example, we completed the parallelization of a Monte Carlo package used in several important codes. Eventually, all applications will be completely parallel-capable. A small number of parallel applications in areas such as global climate modeling and electromagnetic propagation can immediately use the full power of the CS-2 to reach very high levels of performance.

Isotope Separation and Advanced Manufacturing Program

The Laser Program's Isotope Separation and Advanced Manufacturing (ISAM) Program develops and supports computer-based control,

Highlights for 1994

- Began work on the "SuperLab" project, the goal of which is to establish a distributed computing environment spanning LLNL, Los Alamos National Laboratory, and the Sandia National Laboratories in Albuquerque and Livermore.
- Partnered with Los Alamos National Laboratory and an industry group headed by Cray Research in the Industrial Computing Initiative of DOE's High Performance Parallel Processing Project.
- Formed the Center for Computational Sciences and Engineering to develop and apply advanced computational methodologies to solving large-scale scientific and engineering problems arising in DOE energy and environmental and industrial technology mission areas. It is also the organizational home of the High Performance Parallel Processing Project (H4P).
- Began upgrades at ESnet sites to enhance communications capability.
- Upgraded the central computing resource of the National Education Supercomputer Program from a single processor Cray-XMP to a four-processor Cray-YMP-EL.
- Began limited production use of the Meiko-CS2 massively parallel computer, integrated it into the classified supercomputer center supporting defense programs, and began working with users to develop a new generation of parallel-capable applications.
- Formed a CRADA with IT Corporation, a large environmental contractor, to develop a massively parallel simulator of groundwater flow and contaminant migration.

diagnostics, information management, and communications systems for the Uranium Atomic Vapor Laser Isotope Separation (U-AVLIS) effort. ISAM has developed approximately 100 integrated applications in the areas of diagnostics, wavefront control, separator controls and diagnostics, data analysis, data distribution and management, and human–machine interfaces.

U-AVLIS "proof-of-principle" runs are being conducted as part of a plan that will lead to a commercial manufacturing plant. The group is preparing for a major 200-hr run of the separator and a major physics experiment. In the stimulated Raman scattering experiment, the interaction of uranium ions and the laser beam causes changes in the frequency of the scattered beam that conveys structural information.

Joel Buenaventura, a participant in the Science and Engineering Research Semester (SERS), is working with a graphical browser developed for the Human Genome Project.

For the Laser Program's Inertial Confinement Fusion (ICF) Program, Computation Organization code developers perform compiler and application code development, user services, and program maintenance functions for the users of LASNEX (a computer program for modeling the physics of ICF targets). Over the past year, we have been moving this extremely large, computationally intensive program from Cray to workstation platforms in order to significantly reduce run costs and make the code available to a wider audience.

Advanced Scientific Visualization

This year, we developed techniques for understanding the flow fields that are typical in many simulations (such as winds or the propagation of contaminants in groundwater). These new techniques maximize the use of advanced graphical workstations to achieve highly interactive results. For global climate modeling, we have applied volume rendering (for showing a 3D volume in a 2D display) to represent the interaction between winds and clouds, and we have other techniques that represent the flow at specific areas, such as around or through surfaces. Our simulation techniques suit a broad range of applications, including groundwater studies, fluid-flow simulations, and interstellar cloud formations, to name a few. The techniques are being integrated into standard commercial visualization systems to make them available to a greater variety of users.

Applied Computations

The Laboratory is developing advanced computational methodologies to solve large-scale scientific and engineering problems for specific applications in DOE mission areas involving energy, environmental, and industrial technology. Although a multidisciplinary approach is fundamental in this research, we rely heavily upon the strength of our long-standing applied mathematical sciences research program, which emphasizes basic numerical research in computational fluid dynamics, flow in porous media, particle and chemical kinetics, plasma modeling, and general numerical methods for solving large-scale differential/algebraic systems. This algorithmic technology is complemented by the development of advanced software strategies for exploiting state-of-the-art computational hardware, including high-end visualization systems, workstation clusters, and massively parallel computers.

Energy Sciences Network

The Energy Sciences Network (ESnet) is a nationwide communications network for disseminating computer data to support multiple-program, open scientific research. Managed and funded by the DOE's Energy Research (ER) Office of Scientific Computing, ESnet is intended to facilitate remote access to major scientific facilities and resources; provide information dissemination among scientific collaborators throughout all ER programs; support remote experimentation, distributed computing, and collaborative environments; provide widespread access to existing ER supercomputer facilities; and support technology evaluation and prototyping for the technologies that will be important to the development and implementation of the emerging

National Information Infrastructure, the Clinton Administration's initiative to promote the "information superhighway."

The Human Genome Project

The Laboratory's role in the Human Genome Project is to develop efficient biological and computational techniques for chromosome mapping. Our work has become a cornerstone of the entire project, accentuating the importance of software engineering in biomedical research. This year, one task was to automatically integrate more than 10 types of genome physical mapping data. We solved this problem, which had never been attempted on this scale, by turning it into an optimization problem using simulated annealing (a technique for finding a global near-minimum without getting trapped in local minima, analogous to annealing molten glass.) Integrating a 1200-×-300-object physical map on a single workstation took about four days; since we modified the algorithm to run in parallel on a network of 45+ workstations, it can process a 3000-×-750-object physical map in four to six hours. We rewrote our graphical browser tool to handle new types of genome mapping data and to effortlessly link the different modes of map display. We are also providing increased consulting on genomic databases and algorithms to the human- and plant-genome communities.

The Computer Security Technology Center

National and international computer networks are subject to ever-increasing numbers of attacks, attacks recently aided by new automated tools that steal passwords or search for vulnerable computers. The Computer Security Technology Center (CSTC) conducts research and develops innovative ways to prevent harm to computer systems. For example, a tool called the Security Profile Inspector assesses the security of UNIX- and VMS-based systems, reporting system configuration vulnerabilities, bad passwords, and system file integrity violations. Its menus and help features spare system managers the need for extensive system security training.

For Ethernet-based networks, we are developing the Network Intrusion Detector to identify and analyze suspicious traffic and connections. If an incident is suspected, the evidence collected can be shown to site security personnel, law enforcement, or the DOE Computer Incident Advisory Capability (CIAC). CIAC is perhaps the most visible and most critical effort within the CSTC. This team is responsible for helping during computer incidents anywhere within the DOE, and often helps other individuals and organizations around the world. CIAC provides operational support and is the single point of contact for incidents such as computer break-ins, computer viruses, and widespread automated network attacks. This team gathers fast-breaking vulnerability and threat information and disseminates it to the DOE community and elsewhere.

The Department of Energy faces the daunting task of reviewing some 4 billion pages currently in DOE archival storage to determine their suitability for release to the public. CSTC has been very active in this truly challenging project, called the Text Analysis Project. Documents will be read and searched for signs of sensitive information. By automating the process of review, through parsing, root-word analysis, parts-of-speech detection, and similar techniques, we will provide a secondary check of human reviewers to increase assurances that no classified or other sensitive information is accidentally released.

Electronic Commerce

We continued to work with Wright-Patterson Air Force Base to refine the Electronic Commerce project for paperless procurement. The system, in full production this past year, processed over 75% of all commodity procurement actions for an estimated savings of $1.5 million. We received inquiries regarding the design and implementation of this system from within various agencies of the U.S. government and from

LLNL's National Energy Research Supercomputer Center (NERSC) provides high-end computing services to the national community under Energy Research funding. This system serves 500 users at 200 national and international institutions.

Canada and Australia. We plan to adapt Electronic Commerce technology to LLNL departments of Finance and Procurement in 1995.

Numerical Tokamak Project

Tokamak reactors of a scale that will produce electrical power for commercial use are larger than today's research tokamaks; the production-scale systems can be analyzed only by computer simulation. We participate in the Numerical Tokamak Project, a national consortium to simulate the turbulent transport of plasma in tokamak fusion experiments. These simulations will predict the scaling behavior of plasma transport in tokamaks so that we can ascertain the ultimate performance of these devices at production scale. Massively parallel computing and high-speed communications have enormously increased the information available for understanding and appraising solutions to this nationally important scientific problem. We have also created an extensive graphics post-processing system designed to enable computer-supported collaborative work. Collaborators who are widely dispersed can interactively and simultaneously create and view scientific visualizations of the results from large turbulence simulations.

A Preprocessor for High-Performance Computing

The Parallel Data Distribution Preprocessor (PDDP) project is a software effort to implement an efficient parallel programming model. The project is meant specifically for distributed-memory machines, to give the programmer a mechanism for dealing with the nonuniform memory access of these machines. For the past two years the project has focused on a variant of the High Performance Fortran model that is becoming popular among the vendors of massively parallel processing machines. The model is a shared memory model and works by using the parallelism in the program and the locality of data to the advantage of the programmer. Data is worked on locally in each of the processors to get maximum memory bandwidth.

PDDP operates on the BBN TC2000 and Meiko CS-2 computers, and is being ported to the Cray T3D, allowing the same source code to execute on all platforms. Results from applications are very encouraging—performance scales well, and programmers appreciate the simpler model.

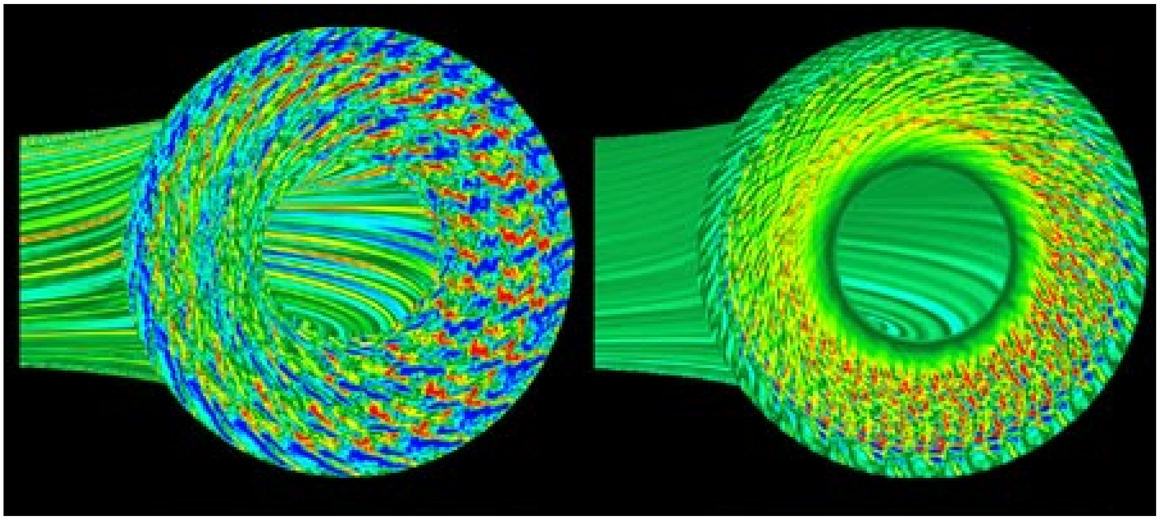
CRADA with Citibank

Although the use of credit cards and electronic funds transfers continues to rise dramatically, the use of paper checks also continues to increase at about 4% per year, and is expected to do so for the next 5–10 years. The handling of checks is costly, error-prone, and time-consuming. There is a clear need to capture checks electronically as early in the processing cycle as possible (i.e., at the first bank the check encounters on its trip through the banking system), to move only electronic bits thereafter, and finally to archive only the electronic images. A key factor is to capture the check with sufficient image quality while minimizing the number of electronic bits. We are in the second year of a three-year CRADA with Citibank on various aspects of check imaging, such as research in Interbank Check Image Exchange and Truncation.

As a consequence of our association with Citibank, we have been instrumental in forming a consortium of major U.S. banks and other financial service providers and equipment vendors. The Financial Services Technology Consortium (FSTC) is conducting a program of precompetitive research, development, and deployment of solutions to problems of privacy, security, reliability, etc., that are encountered when financial services are provided over open public networks. FSTC expects to play a major role in the development of electronic commerce in the U.S.

Open Systems Laboratory

This year, LLNL established the Information Technology Open Systems Laboratory (OSL). The OSL is one of several regionally based sites in the U.S. to provide services on a nationwide scale. It has two comprehensive goals: supporting internal Laboratory programmatic and administrative



These computational models were developed by LLNL researchers for the Numerical Tokamak Project. High-performance computers allow high-resolution visualization of how turbulent mixing of hot core plasma with cool edge plasma prevents fusion ignition (left) and how certain conditions can reduce turbulence (right).

functions, and increasing the effectiveness and efficiency of doing business electronically by promoting the rapid integration of technologies into the Clinton Administration's National Information Infrastructure initiatives.

OSL will promote these goals by working to achieve critical operational goals:

- Develop, commercialize, and deploy computation- and communication-based information systems to facilitate improvements in the financial, education, health care, manufacturing, and government-services sectors of the U.S. economy.
- Provide an information technology testbed and information-systems laboratory for demonstrations, prototyping, and pilot projects.

This laboratory will be used by vendors, systems integrators, and university and national laboratory teams to develop databases, instrumentation systems, communications systems, and software components to model widely distributed information environments. These basic "information superhighway" technologies and applications include functions supporting information analysis, data navigation tools, storage environments, collaboration and telepresence capabilities, and inter-networking and information surety and security.

National Storage Laboratory

High-performance mass storage systems need improvements of two or more orders of magnitude in performance and capacity plus corresponding improvements in architecture and functionality.

LLNL is participating in the National Storage Laboratory (NSL), which was established to investigate, demonstrate, and commercialize new mass storage system architectures to meet such needs. The NSL and closely related projects involve more than 20 participating organizations from industry, DOE, other federal laboratories, universities, and the National Science Foundation supercomputer centers.

Two software development projects are in process: NSL-UniTree and the High Performance Storage System (HPSS). NSL-UniTree features network-attached storage, dynamic storage hierarchies, and capabilities for extensive storage system management. NSL-UniTree has demonstrated more than an order of magnitude improvement in storage-system performance. Release 2 of NSL-UniTree was delivered to IBM for commercialization and is currently used at over a dozen federal agencies. HPSS is focused on developing the next generation of scalable, standards-based, general-purpose, distributable storage systems. It can support storage capacities from gigabytes to petabytes, high-speed data transfers for workstation cluster, and parallel computers through both sequential and scalable parallel storage and retrieval techniques. HPSS is currently in integration testing and planning for its first release. The NSL continues to work on future releases to extend HPSS functionality.

For further information contact C. William McCurdy (510) 422-6383 or J. Joseph Brandt (510) 422-7043.